Joji - 777 : Animated Multi-Character Paintings with a Single Performer

A novel approach to realtime motion capture and choreography

Saad Moosajee
hello@moosajee.co
Independent
USA

Maya Man
mayaman26@gmail.com
Independent
USA

James Bartolozzi
bartolozzij@gmail.com
Independent
USA

ABSTRACT

Tableau paintings, described as "living pictures" often take the form of a painting or photograph in which characters are arranged for picturesque or dramatic effect. These paintings and photographs are inherently defined by the motion of a character, but have traditionally been presented as static two dimensional images. What would it feel like to create a tableau painting that is not static, and is inherently three dimensional? This talk will center around an analysis of the creative and technical work-flows used during the production of Joji - 777, in which a small crew created an animated video where each scene was composed with intention of depicting a moving, 3D tableau painting. The technical complexities of achieving the desired result were amplified due to COVID lockdown restrictions, which at the time of production ruled out the ability to motion capture multiple per-formers at once. Using real time game engine technology in conjunction with skeletal retargeting in post production, a novel approach was developed to allow a single performer to play multiple characters simultaneously.

KEYWORDS

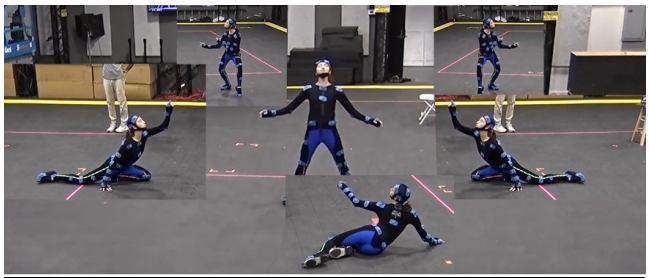
Motion capture, tableau, chiaroscuro, joji, 777

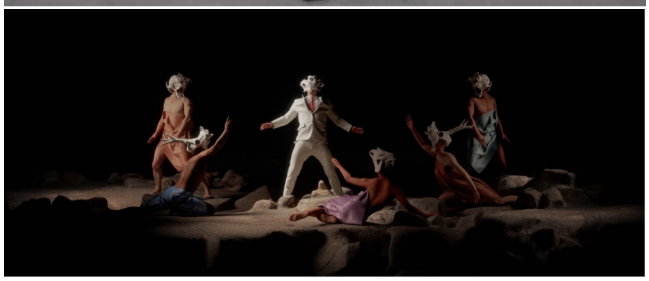
ACM Reference Format:

Saad Moosajee, Maya Man, and James Bartolozzi. 2021. Joji - 777 : Animated Multi-Character Paintings with a Single Performer: A novel approach to realtime motion capture and choreography. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Talks (SIGGRAPH '21 Talks), August 09-13, 2021.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3450623.3469663

1 INTRODUCTION

The music video for 777 was commissioned by record label 88Rising. They were interested in the video ex-pressing a spiritual quality, as the song was written and named after 777, a spiritual number. There was also interest mentioned around integrating a dance component. This combination brought us to tableau paintings from the Renaissance era. Known for their study of the human form, these paintings were interesting reference points visually and anatomically. We concepted a video that would be narrated through different vignettes that echoed tableau paintings, with a twist where each scene would be brought to life through dance. To achieve this, it required the team to meticulously compose each digital character's static pose to match that of a painting, while also having it feel like a believable dance sequence when the characters eventually moved.





2 PAINTING WITH LIGHT

To properly articulate the aesthetic in 3D, we studied in depth how classical painters approached light in their work. The most direct example was Chiaroscuro, which employs high contrast to shape and represent the form of its subject. We initially tested using standard light set ups within a physically based engine, but found the result to be too uniform. We proceeded to develop an adjusted technique within Arnold Render that leveraged its light decay filter to limit the influence and falloff of an individual light object. By assigning a unique light decay filter to each 3D light, we could control exactly how much of the character's body was illuminated. The tradeoff was that this deliberately limited the illumination, which meant each scene required a large volume of lights before it could be fully visible. We found this process to be more akin to painting than to standard 3D lighting. Each light was treated as an object that illuminated or shaped a part of the character's body, rather than as a standard "area" or "fill" that would cover a given radius. It also required we animate hundreds of light objects to preserve the effect in motion.

SIGGRAPH '21 Talks, August 09-13, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8373-8/21/08.
https://doi.org/10.1145/3450623.3469663

Figure 1: A character from the film lit through an array of lights implementing light decay.

3 CHOREOGRAPHY

The choreographer and director worked together to create unique movement vocabularies for each set of characters. This was necessary in order to visually differentiate between the motion styles for the Tableau painting characters and the main character. Working under the constraints of the pandemic, the dancer sent videos to the director that she filmed at home in her apartment. The director would note what movements matched his vision for the characters and then the dancer would take that feedback and continue to choreograph a specific "syntax" for each storyline. By defining clear movement qualities for each set of characters before the shoot and pre-visualizing some of the scenes, the dancer was able to clearly imagine each scenario she performed and focus on embodying the specific energy of each character through her recorded movement. Thus, as a single performer, she trained herself to easily switch between multiple characters' movement.

Once each character's movement throughout the song had been choreographed, the dancer performed at Silver Spoon, a motion capture studio in Brooklyn. There were two shoot dates with two weeks in between each. The first shoot consisted of solely capturing movement data without real time interaction. Then, during the second shoot, the team imported that movement from the first shoot allowing the dancer to react in real time to movement she had performed weeks ago. Watching previous motion data unfold via Unreal Engine on a screen in the studio space, the team was able to map specific markers in the room to where other characters would eventually exist in the final rendered space. This helped the team choreograph more natural looking interaction between characters with a single performer.

4 MOTION CAPTURE PIPELINE

Through real-time compositing of the various motion capture performances in Unreal Engine, the team was able to expand the vocabulary of movement in the production and create believable group behavior. Rather than a dancer representing a different individual, each dancer would represent a different take of the performer Maya Man, that would then be layered into Unreal Engine. In some scenes this meant filming more than 10 unique takes of Maya to map to the different dancers. Once the real time composite was built, the animation team would retarget the dancers to a variety of custom rigged digital humans. These assets were based on real 3D scans, they shared similar skeletal positions and had unique skin shaders which was key in establishing a feeling of individuality.

Although our crew was small and limited by quarantine restrictions, this pipeline enabled us to do many iterations and achieve a photorealistic film level result across the characters in the video. We found that leveraging the real time capabilities of a game engine in the choreography and motion capture process actually allowed for a more fluid and iterative working process than traditional pipelines.

5 SCALING TO CROWDS

After achieving control of movement and lighting within the aesthetic reference, our final challenge was scaling the visual aesthetic to the crowd level to more authentically capture the breadth of characters typically featured in a tableau painting. Given the complexity of the desired character look, we had to overcome challenges for both artist interactivity and render-time performance. High resolution human body scans and textures, intricately detailed character rigs, and accessories like clothing, simulated cloth, and angel wings, all proved to be hurdles in the development process for the crowds pipeline.

Since the choreography in the video was so meticulously designed, we wanted to be able to control the crowd movement in a less-procedural way. Artist tools were built in Houdini to manipulate animation clip transitions with an intuitive and fast user interface. Using these tools, we could create 'wave'-like effects with the crowd animation clip transitions.

Render-time and viewport performance were substantially improved using a variety of level-of-detail and cull-ing techniques. Poly-reducing based on rig-joint proximity maintained detail in the most deformed parts of the character mesh. Independent level of detail controls for layers like helmets and clothes allowed to maintain the quality of the accessory assets independently from the character mesh. Frustum, offscreen, and occlusion culling techniques were all accelerated using character rig points instead of instance point locations or mesh geometry. All characters and accessory meshes had procedurally decimated meshes for three different levels-of-detail. These were interchanged dynamically based on the crowd character's distance to the camera.